

A COUNTER EXAMPLE TO CENTRAL LIMIT THEOREM IN HILBERT SPACES UNDER A STRONG MIXING CONDITION

DAVIDE GIRAUDO AND DALIBOR VOLNÝ

ABSTRACT. We show that in a separable infinite dimensional Hilbert space, uniform integrability of the square of the norm of normalized partial sums of a strictly stationary sequence, together with a strong mixing condition, does not guarantee the central limit theorem.

Keywords: Central limit theorem ; Hilbert space ; mixing conditions ; strictly stationary process

AMS MSC 2010: 60F05 ; 60G10

1. INTRODUCTION AND NOTATIONS

Let $(\Omega, \mathcal{F}, \mu)$ be a probability space and (S, d) a separable metric space. We say that the sequence of random variables $(X_n)_{n \in \mathbb{Z}}$ from Ω to S is *strictly stationary* if for all integer d and all integer k , the d -uple (X_1, \dots, X_d) has the same law as $(X_{k+1}, \dots, X_{k+d})$.

Rosenblatt introduced in [18] the measure of dependence between two sub- σ -algebras \mathcal{A} and \mathcal{B} :

$$\alpha(\mathcal{A}, \mathcal{B}) := \sup \{ |\mu(A \cap B) - \mu(A)\mu(B)|, A \in \mathcal{A}, B \in \mathcal{B} \}.$$

Another one is β -mixing, which is defined by

$$\beta(\mathcal{A}, \mathcal{B}) := \frac{1}{2} \sup \sum_{i=1}^I \sum_{j=1}^J |\mu(A_i \cap B_j) - \mu(A_i)\mu(B_j)|,$$

where the supremum is taken over the finite partitions $\{A_1, \dots, A_I\}$ and $\{B_1, \dots, B_J\}$ of Ω , which consist respectively of elements of \mathcal{A} and \mathcal{B} . It was introduced by Volkonskii and Rozanov in [21].

In order to measure dependence of a sequence of random variables, say $X := (X_j)_{j \in \mathbb{Z}}$ (assumed strictly stationary for simplicity), we define \mathcal{F}_m^n as the σ -algebra generated by the X_j for $m \leq j \leq n$, where $-\infty \leq m \leq n \leq +\infty$.

Then mixing coefficients are defined by

$$(1) \quad \alpha_X(n) := \alpha(\mathcal{F}_{-\infty}^0, \mathcal{F}_n^{+\infty})$$

$$(2) \quad \beta_X(n) := \beta(\mathcal{F}_{-\infty}^0, \mathcal{F}_n^{+\infty}),$$

which will be simply written $\alpha(n)$ (respectively $\beta(n)$) when there is no ambiguity.

We say that the strictly stationary sequence $(X_j)_j$ is α -mixing (respectively β -mixing) if $\lim_{n \rightarrow \infty} \alpha(n) = 0$ (respectively $\lim_{n \rightarrow \infty} \beta(n) = 0$). Sequences which are α -mixing are also called *strong-mixing*. Notice that the inequality $2\alpha(\mathcal{A}, \mathcal{B}) \leq \beta(\mathcal{A}, \mathcal{B})$ for any two sub- σ -algebras \mathcal{A} and \mathcal{B} implies that each β -mixing sequence is strong mixing. We refer the reader to Bradley's book [4] for further information about mixing conditions.

Date: September 2, 2014.

Let $(V, \|\cdot\|)$ be a separable normed space. We can represent a strictly stationary sequence $(X_j)_j$ by $X_j = f \circ T^j$, where $T: \Omega \rightarrow \Omega$ is measurable and measure preserving, that is, $\mu(T^{-1}(S)) = \mu(S)$ for all $S \in \mathcal{F}$ (see [8], p.456, second paragraph).

Given an integer N , we define $S_N(f) := \sum_{j=0}^{N-1} f \circ T^j$ and $(\sigma_N(f))^2 := \mathbb{E} [\|S_N(f)\|^2]$.

When $V = \mathbb{R}^d$, $d \in \mathbb{N}^*$ it is well-known that if $(f \circ T^j)_{j \geq 0}$ satisfies the following assumptions:

- (1) $\lim_{N \rightarrow +\infty} \sigma_N(f) = +\infty$;
- (2) $\int f d\mu = 0$;
- (3) $\lim_{n \rightarrow +\infty} \alpha(n) = 0$;
- (4) the family $\left\{ \frac{\|S_N(f)\|^2}{(\sigma_N(f))^2}, N \geq 1 \right\}$ is uniformly integrable,

then $\left(\frac{1}{\sigma_N(f)} S_N(f) \right)_{N \geq 1}$ converges in distribution to a Gaussian law. It was established for $d = 1$ by Denker [7], Mori and Yoshihara [14] using a blocking argument. Volný [22] gave a proof for d arbitrary based on approximation by an array of independent random variables.

A natural question would be: what if we replace \mathbb{R}^d by another normed space?

First, we restrict ourselves to separable normed spaces in order to avoid measurability issues of sums of random variables. Corollary 10.9. in [11] asserts that a separable Banach space B with norm $\|\cdot\|_B$ is isomorphic to a Hilbert space if and only if for all random variables X with values in B , the conditions $\mathbb{E}[\mathbf{X}] = 0$ and $\mathbb{E}[\|\mathbf{X}\|_B^2] < \infty$ are necessary and sufficient for X to satisfy the central limit theorem. By " \mathbf{X} satisfies the CLT", we mean that if $(\mathbf{X}_j)_{j \geq 1}$ is a sequence of independent random variables, with the same law as X , the sequence $\left(n^{-1/2} \sum_{j=1}^n \mathbf{X}_j \right)_{n \geq 1}$ weakly converges in B . Hence we cannot expect a generalization in a class larger than separable Hilbert spaces. Such a space is necessarily isomorphic to $\mathcal{H} := \ell^2(\mathbb{R})$, the space of square sumable sequences $(x_n)_{n \geq 1}$ endowed with the inner product $\langle x, y \rangle_{\mathcal{H}} := \sum_{n=1}^{+\infty} x_n y_n$. We shall denote by \mathbf{e}_n the sequence whose all terms are 0, except the n -th which is 1. Bold letters denote both random variables taking their values in \mathcal{H} and elements of this space.

General considerations about probability measures and central limit theorem in Banach spaces are contained in Araujo and Giné's book [2].

Notation 1. If $(a_n)_{n \geq 1}, (b_n)_{n \geq 1}$ are sequences of non-negative real numbers, $a_n \lesssim b_n$ means that $a_n \leq C b_n$, where C does not depend on n . In an analogous way, we define $a_n \gtrsim b_n$. When $a_n \lesssim b_n \lesssim a_n$, we simply write $a_n \asymp b_n$.

Our main results are

Theorem A. *There exists a probability space $(\Omega, \mathcal{F}, \mu)$ such that given $0 < q < 1$, we can construct a strictly stationary sequence $\mathbf{X} = (\mathbf{f} \circ T^k) = (\mathbf{X}_k)_{k \in \mathbb{N}}$ defined on Ω , taking its values in \mathcal{H} , such that:*

- a) $\mathbb{E}[\mathbf{f}] = 0$, $\mathbb{E}[\|\mathbf{f}\|_{\mathcal{H}}^p]$ is finite for each p ;
- b) the limit $\lim_{N \rightarrow \infty} \sigma_N(\mathbf{f})$ is infinite;
- c) the process $(\mathbf{X}_k)_{k \in \mathbb{N}}$ is β -mixing, more precisely, $\beta_{\mathbf{X}}(j) = O\left(\frac{1}{j^q}\right)$;

- d) the family $\left\{ \frac{\|S_N(\mathbf{f})\|_{\mathcal{H}}^2}{\sigma_N^2(\mathbf{f})}, N \geq 1 \right\}$ is uniformly integrable;
- e) if $I \subset \mathbb{N}$ is infinite, the family $\left\{ \frac{S_N(\mathbf{f})}{\sigma_N(\mathbf{f})}, N \in I \right\}$ is not tight in \mathcal{H} ; furthermore, given a sequence $(c_N)_{N \geq 1}$ of real numbers going to infinity, we have either
- $\lim_{N \rightarrow +\infty} \frac{\sigma_N(\mathbf{f})}{c_N} = 0$, hence $\left(\frac{S_N(\mathbf{f})}{c_N} \right)_{N \geq 1}$ converges to $\mathbf{0}_{\mathcal{H}}$ in distribution, or
 - $\limsup_{N \rightarrow +\infty} \frac{\sigma_N(\mathbf{f})}{c_N} > 0$, and in this case the collection $\left\{ \frac{S_N(\mathbf{f})}{c_N}, N \geq 1 \right\}$ is not tight.

Theorem A'. Let $(b_N)_{N \geq 1}$ and $(h_N)_{N \geq 1}$ be sequences of positive real numbers, with $\lim_{N \rightarrow \infty} b_N = 0$ and $\lim_{N \rightarrow \infty} h_N = \infty$. Then there exists a strictly stationary sequence $\mathbf{X} := (\mathbf{f} \circ T^k)_{k \in \mathbb{N}} = (\mathbf{X}_k)_{k \in \mathbb{N}}$ of random variables with values in \mathcal{H} such that A, A, A of Theorem A and the following two properties hold:

- b') we have $\sigma_N^2(\mathbf{f}) \lesssim N \cdot h_N$ and $\frac{\sigma_N^2(\mathbf{f})}{N} \rightarrow \infty$;
- c') the process $(\mathbf{X}_k)_{k \in \mathbb{N}}$ is β -mixing, and there is an increasing sequence $(n_k)_{k \geq 1}$ of integers such that for each k , $\beta_{\mathbf{X}}(n_k) \leq b_{n_k}$.

Remark 2. Theorem A shows that Denker's result does not remain valid in its full generality in the context of Hilbert space valued random variables.

Furthermore, a careful analysis of the proof of Proposition 6 shows that for the construction given in Theorem A, we have $\sigma_N^2(f) = N \cdot h(N)$ with h slowly varying in the strong sense. Theorem 1 of [12] does not remain valid in the Hilbert space setting. Indeed, the arguments given in pages 654-655 show that the conditions of Denker's theorem together with the assumption that $\sigma_N^2 = N \cdot h(N)$ with h slowly varying in the strong sense imply those of Theorem 1. These arguments are still true in the Hilbert space setting.

Remark 3. Theorem A' gives a control of the mixing coefficients on a subsequence. When $b_N := N^{-2}$ for example, the construction gives a better estimation for the considered subsequence than what we get by Theorem A.

Tone has established in [20] a central limit theorem for strictly stationary random fields with values in \mathcal{H} under ρ' -mixing conditions. For sequences, these coefficients are defined by

$$\rho'(n) := \sup \left\{ \frac{|\mathbb{E}[\langle \mathbf{f}, \mathbf{g} \rangle_{\mathcal{H}}] - \langle \mathbb{E}[\mathbf{f}], \mathbb{E}[\mathbf{g}] \rangle_{\mathcal{H}}|}{\|\mathbf{f}\|_{\mathbb{L}^2(\mathcal{H})} \|\mathbf{g}\|_{\mathbb{L}^2(\mathcal{H})}} \right\},$$

where the supremum is taken over all the non-zero functions \mathbf{f} and \mathbf{g} such that \mathbf{f} and \mathbf{g} are respectively $\sigma(X_j, j \in S_1)$ and $\sigma(X_j, j \in S_2)$ -measurable, where S_1 and S_2 are such that $\min_{s \in S_1, t \in S_2} |s - t| \geq n$, while $\mathbb{L}^2(\mathcal{H})$ denote the collection of equivalence classes of random variables $\mathbf{X}: \Omega \rightarrow \mathcal{H}$ such that $\|\mathbf{X}\|_{\mathcal{H}}^2$ is integrable.

So "interlaced index sets" can be considered, which is not the case for α and β -mixing coefficient. Taking f and g as characteristic functions of elements of $\mathcal{F}_{-\infty}^0$ and $\mathcal{F}_n^{+\infty}$ respectively, one can see that $\alpha(n) \leq \rho'(n)$, hence ρ' -mixing condition is more restrictive than α -mixing condition.

A partial generalization of the finite dimensional result was proved by Politis and Romano [15], namely, the conditions $\mathbb{E} \|\mathbf{X}_1\|_{\mathcal{H}}^{2+\delta}$ finite for some positive δ and $\sum_j \alpha_{\mathbf{X}}(j)^{\frac{\delta}{2+\delta}}$ guarantees the convergence of $n^{-1/2} \sum_{j=1}^n \mathbf{X}_j$ to a Gaussian random

variable \mathcal{N} , whose covariance operator S satisfies

$$\mathbb{E} [\langle \mathcal{N}, h \rangle^2] = \langle Sh, h \rangle_{\mathcal{H}} = \text{Var}(\langle \mathbf{X}_1, h \rangle) + 2 \sum_{i=1}^{+\infty} \text{Cov}(\langle \mathbf{X}_1, h \rangle, \langle \mathbf{X}_{1+i}, h \rangle).$$

Similar results were obtained by Dehling [6].

Rio's inequality [16] asserts that given two real valued random variables X and Y with finite two order moments,

$$|\mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]| \leq 2 \int_0^{\alpha(\sigma(X), \sigma(Y))} Q_X(u)Q_Y(u)du.$$

It was extended by Merlevède et al. [13], namely, if \mathbf{X} and \mathbf{Y} are two random variables with values in \mathcal{H} , with respective quantile function $Q_{\|\mathbf{X}\|_{\mathcal{H}}}$ and $Q_{\|\mathbf{Y}\|_{\mathcal{H}}}$, then

$$|\mathbb{E}[\langle \mathbf{X}, \mathbf{Y} \rangle_{\mathcal{H}}] - \langle \mathbb{E}[\mathbf{X}], \mathbb{E}[\mathbf{Y}] \rangle_{\mathcal{H}}| \leq 18 \int_0^{\alpha} Q_{\|\mathbf{X}\|_{\mathcal{H}}}(u)Q_{\|\mathbf{Y}\|_{\mathcal{H}}}(u)du,$$

where $\alpha := \alpha(\sigma(\mathbf{X}), \sigma(\mathbf{Y}))$.

From this inequality, they deduce a central limit theorem for a stationary sequence $(\mathbf{X}_j)_{j \in \mathbb{Z}}$ of \mathcal{H} -valued zero-mean random variables satisfying

$$(3) \quad \int_0^1 \alpha^{-1}(u)Q_{\|\mathbf{X}_0\|_{\mathcal{H}}}^2(u)du < \infty,$$

where α^{-1} is the inverse function of $x \mapsto \alpha_{\mathbf{X}}(\lfloor x \rfloor)$.

Discussion after Corollary 1.2 in [17] proves that the later result implies Politis' one.

Relative optimality of condition (3) (cf. [9]) can give a finite-dimensional counter-example to the central limit theorem when this condition is not satisfied. Here, the condition of uniform integrability prevents such counter-examples.

Defining $\alpha_{2, \mathbf{X}}(n) := \sup_{i \geq j \geq n} \alpha(\mathcal{F}_{-\infty}^0, \sigma(\mathbf{X}_i, \mathbf{X}_j))$ and Q_{X_0} the right-continuous inverse of the function $t \mapsto \mu\{\|\mathbf{X}_0\|_{\mathcal{H}} > t\}$ (that is,

$Q_{\mathbf{X}_0}(u) := \inf\{t \in \mathbb{R}, \mu\{\|\mathbf{X}_0\|_{\mathcal{H}} > t\} \leq u\}$), Dedecker and Merlevède have shown in [5] that under the assumption

$$\sum_{k=1}^{+\infty} \int_0^{\alpha_{2, \mathbf{X}}(k)} Q_{\mathbf{X}_0}^2(u)du < \infty,$$

we can find a sequence $(\mathbf{Z}_i)_{i \in \mathbb{N}}$ of Gaussian random variables with values in \mathcal{H} such that almost surely,

$$\left\| \mathbf{S}_n - \sum_{i=1}^n \mathbf{Z}_i \right\|_{\mathcal{H}} = o\left(\sqrt{n \log \log n}\right).$$

2. THE PROOF

2.1. Construction of f . In order to construct a counter-example, we shall need the following lemma, which will be proved later.

We will denote U the Koopman operator associated to T , which acts on measurable functions by $U(f)(x) := f(T(x))$.

Lemma 4. *Let $(u_k)_{k \geq 1} \subset (0, 1)$ be a sequence of numbers. Then there exists a dynamical system $(\Omega, \mathcal{F}, \mu, T)$ and a sequence of random variables $(\xi_k)_{k \geq 1}$ such that*

(1) *for each $k \geq 1$, $\mu(\xi_k = 1) = \mu(\xi_k = -1) = \frac{u_k}{2}$ and $\mu(\xi_k = 0) = 1 - u_k$;*

(2) the random variables $(U^i \xi_k, k \geq 1, i \in \mathbb{Z})$ are mutually independent.

Recall that \mathbf{e}_k is the k -th element of the canonical orthonormal system of $\mathcal{H} = \ell^2(\mathbb{R})$. We define

$$(4) \quad f_k := \sum_{i=0}^{n_k-1} U^{-i} \xi_k \text{ and } \mathbf{f} := \sum_{k=1}^{+\infty} f_k \mathbf{e}_k,$$

where the ξ_i 's are constructed using to Lemma 4 taking $u_k := n_k^{-2}$. Conditions on the increasing sequence of integers $(n_k)_{k \geq 1}$ will be specified latter.

Then $\mathbf{X}_k := \mathbf{f} \circ T^k$ is a strictly stationary sequence. Note that $\|\mathbf{f}\|_{\mathcal{H}}^2$ is an integrable random variable whenever $\sum_k \frac{1}{n_k}$ is convergent. In the sequel, the choice of n_k will guarantee this condition.

2.2. Preliminary results. We express $S_N(f_k)$ as a linear combination of independent random variables. By direct computations, we get

$$(5) \quad f_k = n_k \xi_k + (I - U) \sum_{i=1-n_k}^{-1} (n_k + i) U^i \xi_k,$$

hence

$$S_N(f_k) = n_k \sum_{j=0}^{N-1} U^j \xi_k + \sum_{i=1-n_k}^{-1} (n_k + i) U^i \xi_k - \sum_{i=N-n_k+1}^{N-1} (n_k + i - N) U^i \xi_k.$$

This formula can be simplified if we distinguish the cases $N \geq n_k$ and $n_k < N$ (we break the third sum at the index $i = 0$ if necessary). This gives

$$(6) \quad S_N(f_k) = \sum_{j=0}^{N-1} (N - j) U^j \xi_k + \sum_{j=1-n_k}^{N-n_k} (n_k + j) U^j \xi_k \\ + N \sum_{j=1+N-n_k}^{-1} U^j \xi_k, \quad \text{if } N < n_k,$$

$$(7) \quad S_N(f_k) = n_k \sum_{j=0}^{N-n_k} U^j \xi_k + \sum_{j=N-n_k+1}^{N-1} (N - j) U^j \xi_k \\ + \sum_{j=1-n_k}^{-1} (n_k + j) U^j \xi_k, \quad \text{if } N \geq n_k.$$

The computation of the expectation of the square of partial sums gives

$$(8) \quad \sigma_N^2(f_k) = \begin{cases} \frac{1}{n_k^2} \left(2 \sum_{j=1}^N j^2 + (n_k - N - 1) N^2 \right) & \text{if } N < n_k, \\ \frac{1}{n_k^2} \left(n_k^2 (N - n_k + 1) + 2 \sum_{j=1}^{n_k-1} j^2 \right) & \text{if } N \geq n_k. \end{cases}$$

Notation 5. If N is a positive integer and $(n_k)_{k \geq 1}$ is an increasing sequence of integers, denote by $i(N)$ the unique integer for which $n_{i(N)} \leq N < n_{i(N)+1}$.

Proposition 6. Assume that $(n_k)_{k \geq 1}$ satisfies the condition

(C) there is $p > 1$ such that for each k , $n_{k+1} \geq n_k^p$.

Then $\sigma_N^2(\mathbf{f}) \asymp N \cdot i(N)$.

Proof. Using (8), the fact that $M^3 \asymp \sum_{j=1}^M j^2$ and $\sigma_N^2(\mathbf{f}) = \sum_{k \geq 1} \sigma_N^2(f_k)$, we have

$$(9) \quad \sigma_N^2(\mathbf{f}) \geq \sum_{k=1}^{i(N)} \sigma_N^2(f_k) \asymp N \sum_{j=1}^{i(N)} 1 = N \cdot i(N).$$

From (8) in the case $n_k \geq N$, we deduce

$$(10) \quad \sum_{k \geq i(N)+1} \sigma_N^2(f_k) \lesssim \sum_{k \geq i(N)+1} \frac{N^2}{n_k} \leq \frac{N^2}{n_{i(N)+1}} + \sum_{k \geq i(N)+1} \frac{N^2}{n_k} \frac{1}{n_k^{p-1}}.$$

Since $n_{i(N)+1} \geq N$ and the series $\sum_{k \geq 1} n_k^{1-p}$ is convergent (by the ratio test), we obtain

$$(11) \quad \sum_{k \geq i(N)+1} \sigma_N^2(f_k) \lesssim N + N \sum_{k \geq i(N)+1} \frac{1}{n_k^{p-1}} \lesssim N.$$

Combining (9) and (11), we get

$$(12) \quad N \cdot i(N) \lesssim \sigma_N^2(\mathbf{f}) \lesssim \sum_{k=1}^{i(N)} \sigma_N^2(f_k) + \sum_{k \geq i(N)+1} \sigma_N^2(f_k) \lesssim N \cdot i(N) + N \lesssim N \cdot i(N).$$

□

Proposition 7. Assume that $\sum_k n_k^{-a}$ is convergent for any positive real number a . Then for each integer p , $\|\mathbf{f}\|_{\mathcal{H}}$ has a finite moment of order p .

Proof. We shall use Rosenthal's inequality (Theorem 3, [19]): there exists a constant C depending only on q such that if M is an integer, Y_1, \dots, Y_M are independent real valued zero mean random variables for which $\mathbb{E}|Y_i|^q < \infty$ for each i , then

$$(13) \quad \mathbb{E} \left| \sum_{j=1}^M Y_j \right|^q \leq C \left(\sum_{j=1}^M \mathbb{E}|Y_j|^q + \left(\sum_{j=1}^M \mathbb{E}[Y_j^2] \right)^{q/2} \right).$$

If $q = 2p$ is given then we have

$$(14) \quad \mathbb{E}|f_k|^{2p} \lesssim n_k^{-1} + n_k^{-p} \lesssim n_k^{-1}.$$

□

We provide a sufficient condition for the uniform integrability of the family $\mathcal{S} := \left\{ \frac{\|S_N(\mathbf{f})\|_{\mathcal{H}}^2}{\sigma_N^2(\mathbf{f})}, N \geq 1 \right\}$.

Proposition 8. If $(n_k)_{k \geq 1}$ satisfies (C), then \mathcal{S} is uniformly integrable.

Proof. Let i_0 be such that for each integer k , $n_{k+i_0} \geq n_k^2$. For $N \geq 1$, we have:

$$\frac{\|S_N(\mathbf{f})\|_{\mathcal{H}}^2}{\sigma_N^2(\mathbf{f})} = \sum_{j=1}^{i(N)-1} \frac{|S_N(f_j)|^2}{\sigma_N^2(\mathbf{f})} + \frac{|S_N(f_{i(N)})|^2}{\sigma_N^2(\mathbf{f})} + \frac{|S_N(f_{i(N)+1})|^2}{\sigma_N^2(\mathbf{f})} + \sum_{j \geq i(N)+2} \frac{|S_N(f_j)|^2}{\sigma_N^2(\mathbf{f})},$$

Since $\mathbb{E} \left[\sum_{j=i-i_0}^{i(N)-1} \frac{|S_N(f_j)|^2}{\sigma_N^2(\mathbf{f})} \right] \leq \frac{i_0}{i(N)}$ it is enough to prove that the families

$$\begin{aligned} \mathcal{S}_1 &:= \left\{ \sum_{k=1}^{i(N)-i_0} \frac{|S_N(f_k)|^2}{\sigma_N^2(\mathbf{f})}, N \geq 1 \right\}, \\ \mathcal{S}_2 &:= \left\{ \frac{|S_N(f_{i(N)})|^2}{\sigma_N^2(\mathbf{f})}, N \geq 1 \right\} =: \{u_N, N \geq 1\}, \\ \mathcal{S}_3 &:= \left\{ \frac{|S_N(f_{i(N)+1})|^2}{\sigma_N^2(\mathbf{f})}, N \geq 1 \right\} =: \{v_N, N \geq 1\}, \text{ and} \\ \mathcal{S}_4 &:= \left\{ \sum_{k \geq i(N)+2} \frac{|S_N(f_k)|^2}{\sigma_N^2(\mathbf{f})}, N \geq 1 \right\} \end{aligned}$$

are uniformly integrable. For \mathcal{S}_1 and \mathcal{S}_4 , we shall show that these families are bounded in \mathbb{L}^p for $p \in (1, 2]$ as in (C).

- for \mathcal{S}_1 : using the expression in (7) and (13) with $q := 2p > 2$, we have

$$\begin{aligned} \mathbb{E} [|S_N(f_k)|^{2p}] &\leq C \left(2 \sum_{j=1}^{n_k} \frac{j^{2p}}{n_k^2} + \frac{n_k^{2p}(N - n_k)}{n_k^2} \right) + C \left(2 \sum_{j=1}^{n_k} \frac{j^2}{n_k^2} + \frac{(N - n_k)n_k^2}{n_k^2} \right)^p \\ &\lesssim \frac{1}{n_k^2} \left(n_k^{2p+1} + (N - n_k)n_k^{2p} \right) + \frac{1}{n_k^{2p}} \left(n_k^3 + (N - n_k)n_k^2 \right)^p \\ &= \frac{N n_k^{2p}}{n_k^2} + \frac{N^p n_k^{2p}}{n_k^{2p}} \\ &= N n_k^{2(p-1)} + N^p \end{aligned}$$

hence

$$\|S_N(f_k)^2\|_p \lesssim N^{1/p} n_k^{2\frac{p-1}{p}} + N,$$

which gives

$$\begin{aligned} \left\| \sum_{k=1}^{i(N)-i_0} \frac{|S_N(f_k)|^2}{\sigma_N^2(\mathbf{f})} \right\|_p &\lesssim \frac{\sum_{k=1}^{i(N)-i_0} (N^{1/p} n_k^{2\frac{p-1}{p}} + N)}{\sigma_N^2(\mathbf{f})} \\ &\lesssim \frac{N^{1/p} i(N) n_{i(N)-i_0}^{2\frac{p-1}{p}} + N i(N)}{\sigma_N^2(\mathbf{f})} \\ &\lesssim \left(\frac{n_{i(N)-i_0}^2}{N} \right)^{1-1/p} + 1 \end{aligned}$$

From (9), we get

$$\left\| \sum_{k=1}^{i(N)-1} \frac{|S_N(f_k)|^2}{\sigma_N^2(\mathbf{f})} \right\|_p \lesssim \frac{n_{i(N)}^{2\frac{p-1}{p}}}{n_{i(N)}} + 1 = n_{i(N)}^{\frac{p-2}{p}} + 1.$$

Since $p - 2 \leq 0$, we obtain that \mathcal{S}_1 is bounded in \mathbb{L}^p hence uniformly integrable.

- for \mathcal{S}_2 : using (7) in the case $n_k \leq N$ and Proposition 6, we get

$$(15) \quad \|u_N\|_1 \lesssim \frac{N}{\sigma_N^2(\mathbf{f})} \lesssim \frac{1}{i(N)}.$$

Since $\|u_N\|_1 \rightarrow 0$ and $u_N \in \mathbb{L}^1$ for each N , the family \mathcal{S}_2 is uniformly integrable.

- for \mathcal{S}_3 : using (6) in the case $n_k > N$ and Proposition 6, we get

$$(16) \quad \|v_N\|_1 \lesssim \frac{N^2}{n_{i(N)+1}\sigma_N^2(\mathbf{f})} \lesssim \frac{N}{N \cdot i(N)}.$$

Since $\|v_N\|_1 \rightarrow 0$ and $v_N \in \mathbb{L}^1$ for each N , the family \mathcal{S}_3 is uniformly integrable.

- for \mathcal{S}_4 : as for \mathcal{S}_1 , we shall show that this family is bounded in \mathbb{L}^p with $p \in (1, 2]$. We have, using (6) and (13)

$$\begin{aligned} \mathbb{E} \left[|S_N(f_k)|^{2p} \right] &\lesssim \frac{1}{n_k^2} (N^{2p+1} + N^{2p}(n_k - N)) + \frac{1}{n_k^{2p}} (N^3 + (n_k - N)N^2)^p \\ &= \frac{N^{2p}}{n_k} + \frac{N^{2p}}{n_k^p} \\ &\lesssim \frac{N^{2p}}{n_k} \end{aligned}$$

as $N \leq n_k$. We thus get that

$$\left\| \sum_{k \geq i(N)+2} |S_N(f_k)|^2 \right\|_p \lesssim N^2 \sum_{k \geq i(N)+2} \frac{1}{n_k^{1/p}}.$$

Also, using (8), we have

$$\sigma_N^2(\mathbf{f}) \gtrsim N^2 \sum_{k \geq i(N)+1} \frac{1}{n_k}.$$

The condition $n_{k+1} \geq n_k^p$ gives boundedness in \mathbb{L}^p of \mathcal{S}_4 .

This concludes the proof of A. \square

Proposition 9. Assume that $(n_k)_{k \geq 1}$ is such that \mathcal{S} is uniformly integrable and $\sum_k n_k^{-1}$ is convergent. Then for each $I \subset \mathbb{N}$ infinite, the collection $\left\{ \frac{S_N(\mathbf{f})}{\sigma_N(\mathbf{f})}, N \in I \right\}$ is not tight in \mathcal{H} . Its finite-dimensional distributions converge to 0 in probability.

Furthermore, if $(c_N)_{N \geq 0}$ is a sequence of positive numbers going to infinity, we have either

- $\lim_{N \rightarrow +\infty} \frac{\sigma_N(\mathbf{f})}{c_N} = 0$, hence $\left(\frac{S_N(\mathbf{f})}{c_N} \right)_{N \geq 1}$ converges to $\mathbf{0}_{\mathcal{H}}$ in distribution, or
- $\limsup_{N \rightarrow +\infty} \frac{\sigma_N(\mathbf{f})}{c_N} > 0$, and in this case the sequence $\left\{ \frac{S_N(\mathbf{f})}{c_N}, N \geq 1 \right\}$ is not tight.

Proof. We first prove that the finite dimensional distributions of $\frac{S_N(\mathbf{f})}{\sigma_N(\mathbf{f})}$ converge weakly to 0.

For each $d \in \mathbb{N}$, we have $\frac{\langle S_N(\mathbf{f}), \mathbf{e}_d \rangle_{\mathcal{H}}}{\sigma_N(\mathbf{f})} \rightarrow 0$ in distribution. Indeed, we have by (5) that $\langle S_N(\mathbf{f}), \mathbf{e}_d \rangle_{\mathcal{H}} = n_d \sum_{i=0}^{N-1} U^i \xi_d + (I - U^N) \sum_{i=1-N_d}^{-1} (n_d + i) U^i \xi_d$. We conclude

noticing that $\sigma_N(\mathbf{f})^{-1}(I - U^N) \sum_{i=1-n_d}^{-1} (n_d + i) U^i \xi_d$ goes to 0 in probability as N goes to infinity, using Proposition 6 and the estimate

$$\mathbb{E} \left(n_d \sum_{i=0}^{N-1} U^i \xi_d \right)^2 = N \lesssim \frac{\sigma_N^2(\mathbf{f})}{i(N)}$$

This can be extended replacing \mathbf{e}_d by any $\mathbf{v} \in \mathcal{H}$ by an application of Theorem 4.2. in [3]. By Proposition 4.15 in [2], the only possible limit is the Dirac measure at $\mathbf{0}_{\mathcal{H}}$.

Assume that the sequence $\left\{ \frac{S_N(\mathbf{f})}{\sigma_N(\mathbf{f})}, N \geq 1 \right\}$ is tight. The sequence $\left(\frac{\|S_N(\mathbf{f})\|_{\mathcal{H}}^2}{\sigma_N^2(\mathbf{f})} \right)_{N \geq 1}$ is a uniformly integrable sequence of random variables of mean 1. A weakly convergent subsequence would go to $\mathbf{0}_{\mathcal{H}}$. According to Theorem 5.4 in [3], we should have that the limit random variable has expectation 1. This contradiction gives the result when $I = \mathbb{N} \setminus \{0\}$. Applying this reasoning to subsequences, one can see that for any infinite subset I of $\mathbb{N} \setminus \{0\}$, the family $\left\{ \frac{S_N(\mathbf{f})}{\sigma_N(\mathbf{f})}, N \in I \right\}$ is not tight.

Let $(c_N)_{N \geq 1}$ be a sequence of positive real numbers such that $\lim_{N \rightarrow +\infty} c_N = +\infty$.

- first case: $\frac{\sigma_N(\mathbf{f})}{c_N}$ converges to 0. In this case, the sequence $\left(\frac{\|S_N(\mathbf{f})\|^2}{c_N^2} \right)_{N \geq 1}$ converges to 0 in \mathbb{L}^1 , hence the sequence $\left(\frac{S_N(\mathbf{f})}{c_N} \right)_{N \geq 1}$ converges in distribution to $\mathbf{0}_{\mathcal{H}}$.
- second case: $\limsup_{N \rightarrow \infty} \frac{\sigma_N(\mathbf{f})}{c_N} > 0$. Hence there is some $r > 0$ and a sequence of integers $l_i \uparrow \infty$ such that for each i , $\frac{\sigma_{l_i}(\mathbf{f})}{c_{l_i}} \geq \frac{1}{r}$, that is, $c_{l_i} \leq r \sigma_{l_i}(\mathbf{f})$.

Assume that the family $\left\{ \frac{S_{l_i}(\mathbf{f})}{c_{l_i}}, i \geq 1 \right\}$ is tight. This means that given a positive ε , one can find a compact set $K = K(\varepsilon)$ such that for each i , $\mu \left\{ \frac{S_{l_i}(\mathbf{f})}{c_{l_i}} \in K \right\} > 1 - \varepsilon$. We can assume that this compact set is convex and contains 0 (we consider the closed convex hull of $K \cup \{0\}$, which is compact by Theorem 5.35 in [1]). Then we have

$$\begin{aligned} \left\{ \frac{S_{l_i}(\mathbf{f})}{c_{l_i}} \in K \right\} &= \left\{ \frac{S_{l_i}(\mathbf{f})}{\sigma_{l_i}(\mathbf{f})} \in \frac{c_{l_i}}{\sigma_{l_i}(\mathbf{f})} K \right\} \\ &\subset \left\{ \frac{S_{l_i}(\mathbf{f})}{\sigma_{l_i}(\mathbf{f})} \in rK \right\}, \end{aligned}$$

and we would deduce tightness of $\left\{ \frac{S_{l_i}(\mathbf{f})}{\sigma_{l_i}(\mathbf{f})}, i \geq 1 \right\}$, which cannot happen.

Remark 10. In the second case, it may happen that the finite dimensional distributions does not converge to degenerate ones, for example with $c_N := N$.

□

2.3. Proof of Theorem A. Notice that if $n_{k+1} \geq n_k^p$ for some $p > 1$ and $n_1 = 2$, then $n_k \geq 2^{p^k}$, hence the condition of Proposition 7 is fulfilled. We get A since each f_k has expectation 0.

We denote $[x] := \sup \{k \in \mathbb{Z}, k \leq x\}$ the integer part of the real number x .

Proposition 11. Let $p > 1$. With $n_k := \lfloor 2^{p^k} \rfloor$ (which satisfies (C)), we have for each positive integer l ,

$$\beta_{\mathbf{X}}(l) \lesssim \frac{1}{l^{\frac{1}{p}}}.$$

Proof. We define $\beta_k(n)$ as the n -th β -mixing coefficient of the sequence $(f_k \circ T^i)_{i \geq 0}$.

By Lemma 5 of [10], we have the estimate $\beta_k(0) \leq 4n_k^{-1}$ for each k . Using then Proposition 4 of this paper (cf. [4] for a proof), we get that $\beta_{\mathbf{X}}(n_k) \lesssim \sum_{j \geq k} \frac{1}{n_j}$ for each integer k . Since $p^i \geq i$ for i large enough,

$$\sum_{j \geq k} \frac{1}{n_j} = \sum_{i=0}^{+\infty} \frac{1}{2^{p^{i+k}}} = \sum_{i=0}^{+\infty} \frac{1}{2^{p^i p^k}} \lesssim \sum_{i=0}^{+\infty} \frac{1}{2^i} \frac{1}{2^{p^k}} = \frac{2}{2^{p^k}},$$

we get

$$\beta_{\mathbf{X}}(N) \leq \beta_X(n_{i(N)}) \lesssim \frac{1}{n_{i(N)}} = \frac{1}{n_{i(N)+1}^{1/p}} \leq \frac{1}{N^{1/p}}.$$

□

This proves A. For any p , the choice $n_k := \lfloor 2^{p^k} \rfloor$ satisfies the condition of Proposition 8, which proves A. We conclude the proof by Proposition 9.

Remark 12. For each of these choices, $\sigma_N^2(\mathbf{f})$ behaves asymptotically like $N \log \log N$. Theorem A' shows that we can construct a process which satisfies the same asymptotic behavior of partial sums and has a variance close to a linear one.

A question would be: can we construct a strictly stationary sequence with all the properties of Theorem A, except A which is replaced by an assumption of linear variance?

2.4. Proof of Theorem A'. Let $(h_N)_{N \geq 1}$ be the sequence involved in Theorem A'. We define for an integer u the quantity $h^{-1}(u) := \inf \{j \in \mathbb{N}, h_j \geq u\}$.

If $(b_k)_{k \geq 1}$ is the given sequence (that can be assumed decreasing), we define inductively

$$(17) \quad n_{k+1} := \max \left\{ n_k^2, \left\lfloor \frac{2^k}{b_{n_k}} \right\rfloor, h^{-1}(k) \right\}.$$

Let N be an integer. We assume without loss of generality that the growth of the sequence $(h_N)_{N \geq 1}$ is slow enough in order to guarantee that there exists k such that $N = h^{-1}(k)$. We then have $i(N) \leq k+1 \leq h_N + 1$, hence using Proposition 6, we get b').

We have $n_k \geq 2^{2^k}$ hence by a similar argument as in the proof of Theorem A, A is satisfied.

By a similar argument as in [10], we get $\beta_{\mathbf{X}}(n_k) \leq b_{n_k}$, hence c') holds.

Remark 13. By (3), we cannot expect the relationship $\beta_{\mathbf{X}}(\cdot) \leq b$ for the whole sequence.

Since for each k , $n_{k+1} \geq n_k^2$, Proposition 8 and 9 apply. This concludes the proof of Theorem A'.

Proof of Lemma 4. Let $\Omega := [0, 1]^{\mathbb{N}^* \times \mathbb{Z}}$, where $[0, 1]$ is endowed with Borel σ -algebra and Lebesgue measure, and Ω with the product structure.

For $(k, j) \in \mathbb{N}^* \times \mathbb{Z}$ and $S \subset [0, 1]$, let $P_{k,j}(S) := \prod_{(i_1, i_2) \in \mathbb{N}^* \times \mathbb{Z}} S_{i_1, i_2}$, where $S_{i_1, i_2} = S$ if $(i_1, i_2) = (k, j)$ and $[0, 1]$ otherwise. Then we define

$$A_{k,j}^+ := P_{k,j}([0, 2^{-1}(u_k)^{-1}]),$$

$$A_{k,j}^- := P_{k,j}([2^{-1}(u_k)^{-1}, (u_k)^{-1}]),$$

$$A_{k,j}^{(0)} := P_{k,j}([(u_k)^{-1}, 1]),$$

the map T by $T\left((x_{k,j})_{(k,j) \in \mathbb{N}^* \times \mathbb{Z}}\right) := (x_{k,j+1})_{(k,j) \in \mathbb{N}^* \times \mathbb{Z}}$, and

$$\xi_k := \chi_{A_{k,0}^+} - \chi_{A_{k,0}^-}.$$

□

Acknowledgments. The authors would like to thank both referees for helpful comments, and for suggesting Remark 2.

REFERENCES

- [1] C. D. Aliprantis and K. C. Border, *Infinite dimensional analysis*, third ed., Springer, Berlin, 2006, A hitchhiker's guide. MR 2378491 (2008m:46001)
- [2] A. Araujo and E. Giné, *The central limit theorem for real and Banach valued random variables*, John Wiley & Sons, New York-Chichester-Brisbane, 1980, Wiley Series in Probability and Mathematical Statistics. MR 576407 (83e:60003)
- [3] P. Billingsley, *Convergence of probability measures*, John Wiley & Sons Inc., New York, 1968. MR 0233396 (38 #1718)
- [4] R. C. Bradley, *Introduction to strong mixing conditions. Vol. 1*, Kendrick Press, Heber City, UT, 2007. MR 2325294 (2009f:60002a)
- [5] J. Dedecker and F. Merlevède, *On the almost sure invariance principle for stationary sequences of Hilbert-valued random variables*, Dependence in probability, analysis and number theory, Kendrick Press, Heber City, UT, 2010, pp. 157–175. MR 2731073 (2012c:60095)
- [6] H. Dehling, *Limit theorems for sums of weakly dependent Banach space valued random variables*, Z. Wahrsch. Verw. Gebiete **63** (1983), no. 3, 393–432. MR 705631 (85g:60013)
- [7] M. Denker, *Uniform integrability and the central limit theorem for strongly mixing processes*, Dependence in probability and statistics (Oberwolfach, 1985), Progr. Probab. Statist., vol. 11, Birkhäuser Boston, Boston, MA, 1986, pp. 269–289. MR 899993 (88h:60044)
- [8] J. L. Doob, *Stochastic processes*, John Wiley & Sons, Inc., New York; Chapman & Hall, Limited, London, 1953. MR 0058896 (15,445b)
- [9] P. Doukhan, P. Massart, and E. Rio, *The functional central limit theorem for strongly mixing processes*, Ann. Inst. H. Poincaré Probab. Statist. **30** (1994), no. 1, 63–82. MR 1262892 (95b:60030)
- [10] D. Giraud and D. Volný, *A strictly stationary β -mixing process satisfying the central limit theorem but not the weak invariance principle*, Stochastic Processes and their Applications **124** (2014), no. 11, 3769 – 3781.
- [11] M. Ledoux and M. Talagrand, *Probability in Banach spaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 23, Springer-Verlag, Berlin, 1991, Isoperimetry and processes. MR 1102015 (93c:60001)
- [12] F. Merlevède and M. Peligrad, *On the weak invariance principle for stationary sequences under projective criteria*, J. Theoret. Probab. **19** (2006), no. 3, 647–689. MR 2280514 (2007m:60059)
- [13] F. Merlevède, M. Peligrad, and S. Utev, *Sharp conditions for the CLT of linear processes in a Hilbert space*, J. Theoret. Probab. **10** (1997), no. 3, 681–693. MR 1468399 (99e:60022)
- [14] T. Mori and K. Yoshihara, *A note on the central limit theorem for stationary strong-mixing sequences*, Yokohama Math. J. **34** (1986), no. 1-2, 143–146. MR 886062 (88g:60068)
- [15] D. N. Politis and J. P. Romano, *Limit theorems for weakly dependent Hilbert space valued random variables with application to the stationary bootstrap*, Statist. Sinica **4** (1994), no. 2, 461–476. MR 1309424 (95j:60013)
- [16] E. Rio, *Covariance inequalities for strongly mixing processes*, Ann. Inst. H. Poincaré Probab. Statist. **29** (1993), no. 4, 587–597. MR 1251142 (94j:60045)

- [17] E. Rio, *Théorie asymptotique des processus aléatoires faiblement dépendants*, Mathématiques & Applications (Berlin) [Mathematics & Applications], vol. 31, Springer-Verlag, Berlin, 2000. MR 2117923 (2005k:60001)
- [18] M. Rosenblatt, *A central limit theorem and a strong mixing condition*, Proc. Nat. Acad. Sci. U. S. A. **42** (1956), 43–47. MR 0074711 (17,635b)
- [19] H. P. Rosenthal, *On the subspaces of L^p ($p > 2$) spanned by sequences of independent random variables*, Israel J. Math. **8** (1970), 273–303. MR 0271721 (42 #6602)
- [20] C. Tone, *Central limit theorems for Hilbert-space valued random fields satisfying a strong mixing condition*, ALEA Lat. Am. J. Probab. Math. Stat. **8** (2011), 77–94. MR 2754401 (2012c:60069)
- [21] V. A. Volkonskiĭ and Y. A. Rozanov, *Some limit theorems for random functions. I*, Teor. Veroyatnost. i Primenen **4** (1959), 186–207. MR 0105741 (21 #4477)
- [22] D. Volný, *Approximation of stationary processes and the central limit problem*, Probability theory and mathematical statistics (Kyoto, 1986), Lecture Notes in Math., vol. 1299, Springer, Berlin, 1988, pp. 532–540. MR 936028 (89d:60054)